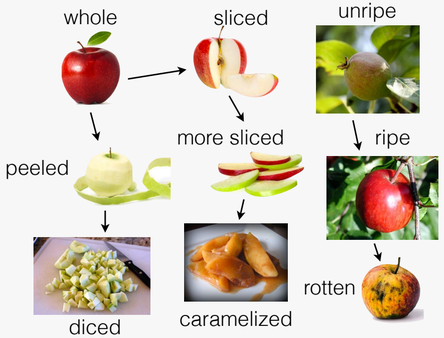
Discovering States and Transformations in Image Collections

Phillip Isola*, Joseph J. Lim*, Edward H. Adelson

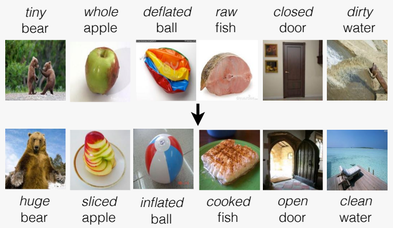
States and transformations

States and transformations describe physical variation within an object class.



Adjectives and antonyms

We define a state as an adjective, and a transformation as a transition from a one adjective to its antonym.



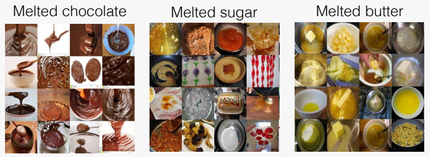
Draped
Loose ⇔ Tight
Heavy ⇔ Lightweight
Crumpled
Crinkled

Unripe ⇔ Ripe
Raw ⇔ Cooked
Burnt
Mashed
Pureed

Sunny ⇔ Foggy
Murky ⇔ Clear
Sunny ⇔ Cloudy
Windblown
Muddy ⇔ Dry

Engraved
Broken
Narrow ⇔ Wide
Ancient ⇔ Modern
Curved ⇔ Straight

Adjectives modify different images in similar ways

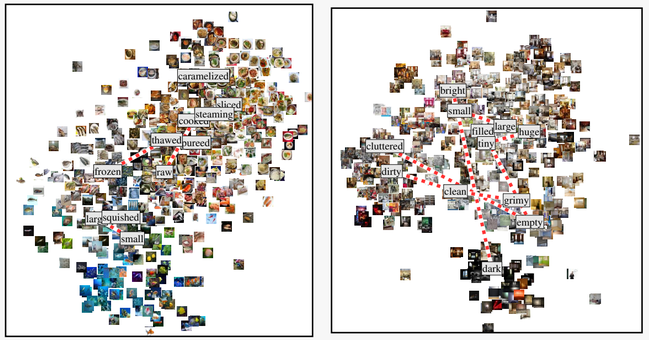


c.f. *Beyond nouns: Exploiting prepositions and comparative adjectives for learning visual classifiers.* Gupta & Davis, ECCV, 2008
Describing objects by their attributes, CVPR 2009
Relative Attributes. Parikh & Grauman, ICCV 2011
Constrained Semi-Supervised Learning using Attributes and Comparative Attributes. Shrivastava et al., ECCV 2012

Dataset

Image collection for each noun, depicting physical state variation that noun can undergo.

249 Nouns
115 Adjectives
70k images

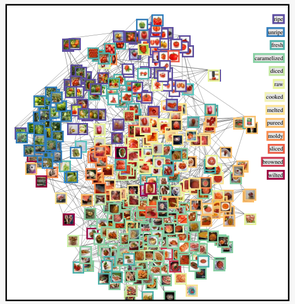


Tasks

Input: *Tomato*

Never before seen image class

1. Identify states



2. Rank images by transformations

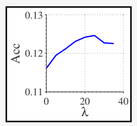
Given a novel image class, we parse the intra-class variation into a set of states and transformations depicted in that photos of that class.

Collection parsing

$$\log p(\mathbf{A}|\mathbf{I}) = \sum_i \phi(A_i|I_i) + \lambda \sum_{i,j \in \mathcal{N}} \psi(A_i, A_j|I_i, I_j) + \log Z$$

$$\psi(A_i, A_j|I_i, I_j) = \mathbb{1}(A_i \neq A_j)\left(\frac{\xi + e^{-\gamma\|f(I_i)-f(I_j)\|^2}}{\xi + 1}\right)$$

$$\arg\min_{\mathcal{P}} \phi(A|I_s) + \phi(\text{ant}(A)|I_t) + \frac{1}{|\mathcal{P}|}\sum_{i=1}^{|\mathcal{P}|}\|\mathbf{f}_{\mathcal{P}_i} - \mathbf{f}_{\mathcal{P}_{i+1}}\|_1$$



Using collection helps

Results

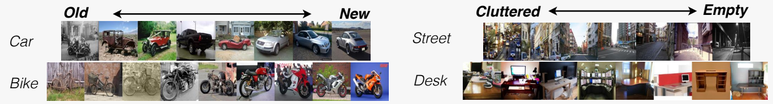
State and transformation discovery

Ranking images according to a transformation

$\phi(A|I) \longrightarrow$ ρ = 0.35

$\phi(A|I) - \phi(\text{ant}(A)|I) \longrightarrow$ ρ = **0.46**

Antonyms pin down a meaningful dimension of variation



How well can the system generalize?